SAT and Hybrid models of the Car-Sequencing problem

Christian Artigues, Emmanuel Hebrard, Valentin Mayer-Eichberger, Mohamed Siala, Toby Walsh

LAAS-CNRS NICTA UNSW AUSTRALIA

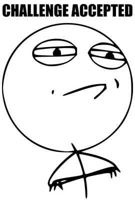
Uppsala, Sweden

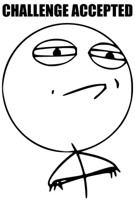
Outline

How did it start?

Mohamed	Valentin
CP lover.. We have Global Constraints!	The SAT-revolution.. Encode Finite domain CSP into SAT !
Hybrid SAT/CP techniques	'advanced' SAT encoding

The Car-sequencing problem





CP/SAT Solving

SAT & CP :

- to encode into SAT or to use global constraints?
- Can we get the best from both approaches?
  →A key concept in hybrid solvers : explaining constraints

An explanation is a set of assignments/prunings triggering a failure/filtering.

example

Cardinality Constraint : $\sum_{i=1}^{n} x_i \leq k$; $D(x_i) = \{0, 1\}$.
$x_i \leftarrow 1$ is pruned if we already have k appearances of the value 1.

$$\{x_j \leftarrow 1 | D(x_j) = \{1\}\} \rightarrow x_i \not\leftarrow 1 .$$

CP/SAT Solving

SAT & CP :

- to encode into SAT or to use global constraints?
- Can we get the best from both approaches?
  →A key concept in hybrid solvers : explaining constraints

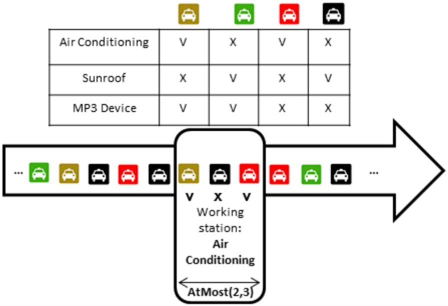
An explanation is a set of assignments/prunings triggering a failure/filtering.

example

Cardinality Constraint : $\sum_{i=1}^{n} x_i \leq k$; $D(x_i) = \{0, 1\}$.
$x_i \leftarrow 1$ is pruned if we already have k appearances of the value 1.

$$\{x_j \leftarrow 1 | D(x_j) = \{1\}\} \quad \rightarrow x_i \not\leftarrow 1 \ .$$

Car-sequencing



Constraints

- Each class c is associated with a demand D_c.
- For each option j, each sub-sequence of size q_j must contain at most u_j cars requiring the option j.

Modelling in CP

Variables :

- n integer variables $\{x_1, \dots, x_n\}$ taking values in $\{1, \dots, k\}$
- nm Boolean variables $\{y_1^1, \dots, y_n^m\}$

Constraints:

1. *Demand constraints* : for each class $c \in \{1..k\}$

$$|\{i \mid x_i = c\}| = D_c^{class}.$$

$$\rightarrow \textsc{Gcc}$$

2. *Capacity constraints* : for each option $j \in \{1..m\}$, for each slot $i \in \{1, \dots, n - q_j + 1\}$.

$$\sum_{l=i}^{i+q_j-1} y_l^j \leq u_j.$$

$\rightarrow$ Gsc, AtMostSeqCard or AtMostSeqCard$\oplus$ Gsc

.

Modelling in CP

Variables :

- n integer variables $\{x_1, \ldots, x_n\}$ taking values in $\{1, \ldots, k\}$
- nm Boolean variables $\{y_1^1, \ldots, y_n^m\}$

Constraints:

1. *Demand constraints* : for each class $c \in \{1..k\}$
$$|\{i \mid x_i = c\}| = D_c^{class}.$$
$\rightarrow$ Gcc
2. *Capacity constraints* : for each option $j \in \{1..m\}$, for each slot $i \in \{1, \ldots, n - q_j + 1\}$.
$$\sum_{l=i}^{i+q_j-1} y_l^j \leq u_j.$$
$\rightarrow$ Gsc, AtMostSeqCard or AtMostSeqCard$\oplus$ Gsc

.

Definition

Definition

$\textsc{AtMostSeqCard}(u, q, d, [x_1, \ldots, x_n]) \Leftrightarrow$

$$\bigwedge_{i=0}^{n-q} \left(\sum_{l=1}^{q} x_{i+l} \leq u\right) \wedge \left(\sum_{i=1}^{n} x_i = d\right)$$

Definition

Definition

$\textsc{AtMostSeqCard}(u, q, d, [x_1, \ldots, x_n]) \Leftrightarrow$

$$\bigwedge_{i=0}^{n-q} \left(\sum_{l=1}^{q} x_{i+l} \leq u\right) \wedge \left(\sum_{i=1}^{n} x_i = d\right)$$

Example $\textsc{AtMostSeqCard}(2, 4, 4, [x_1, \ldots, x_7])$

0 1 1 0 1 1 0

1 1 0 0 1 0 1

The propagator

- leftmost: computes an assignment w maximizing the cardinality of the sequence with respect to the AtMost constraints.
- Let $max(i)$ be the maximum cardinality of the q subsequences involving x_i when computing leftmost[i].

0 1 0 . 0 1 0

$cardinality = 1$

$cardinality = 0$

$cardinality = 1$

$max(4) = 1$

- $Left[i] = \sum_{j=1}^{j=i} \texttt{leftmost}[j]$.
- $Right[i]$: same as *Left* but in the reverse sense, i.e. $[x_n, .., x_1]$.
- Example : with AtMost(2,5):

$\mathcal{D}(x_i)$	0	.	.	.	.	1	.	.	
$max(i)$	0	1	2	2	2	2	1	2	2
$\texttt{leftmost}[i]$	0	1	0	0	0	1	1	0	0
$Left[i]$	0	1	1	1	1	1	2	2	2

Domain consistency

- DC on each AtMost: $(\sum_{l=1}^{q} x_{i+l} \leq u)$
- DC on $\sum_{i=1}^{n} x_i = d$
- If $Left[n] < d$ Then *fail*
- If $Left[n] = d$ and $Left[i] + Right[n-i+1] \leq d$ Then $\mathcal{D}(x_i) \leftarrow \{0\}$
- If $Left[n] = d$ and $Left[i-1] + Right[n-i] < d$ Then $\mathcal{D}(x_i) \leftarrow \{1\}$

Explaining AtMostSeqCard: the key idea

Explaining Failure

1. If a failure is triggered by a cardinality constraint (i.e. $(\sum_{l=1}^{q} x_{i+l} \leq u)$ or $\sum_{i=1}^{n} x_i = d$), then it is easy to generate an explanation.
2. If a failure triggered by $Left[n] < d$, a naive explanation would be the set of all assignments in the sequence.

Some observations

Let S : 1 1 0 0 . subject to ATMOST(2/5).
→`leftmost` on S gives 1 1 0 0 0

Consider the sequence S_0 : 1 1 . 0 .
→`leftmost` on S_0 gives 1 1 0 0 0

Some observations

Let S : 1 1 0 0 . subject to AtMost(2/5).
→`leftmost` on S gives 1 1 0 0 0

Consider the sequence S_0 : 1 1 . 0 .
→`leftmost` on S_0 gives 1 1 0 0 0

$$\{x_i \leftarrow 0 \mid max(i) = u\}$$

Some observations

Let S : 1 1 0 0 . subject to AtMost(2/5).
→`leftmost` on S gives 1 1 0 0 0

Consider the sequence S_0 : 1 1 . 0 .
→`leftmost` on S_0 gives 1 1 0 0 0

$$\{x_i \leftarrow 0 \mid max(i) = u\}$$

Consider the sequence S_2 : . 1 0 0 .
→`leftmost` on S_2 gives 1 1 0 0 0

Some observations

Let S : 1 1 0 0 . subject to AtMost(2/5).
→`leftmost` on S gives 1 1 0 0 0

Consider the sequence S_0 : 1 1 . 0 .
→`leftmost` on S_0 gives 1 1 0 0 0

$$\{x_i \leftarrow 0 \mid max(i) = u\}$$

Consider the sequence S_2 : . 1 0 0 .
→`leftmost` on S_2 gives 1 1 0 0 0

$$\{x_i \leftarrow 1 \mid max(i) \neq u\}$$

Theorem

Theorem

Let S be the set of all assignments,
$S^{\cdot *} = S \setminus (\{x_i \leftarrow 0 \mid max(i) = u\} \cup \{x_i \leftarrow 1 \mid max(i) \neq u\})$*, then*
$S^{\cdot *}$ *is a valid explanation.*

→runs in $O(n)$ since we call `leftmost` once.

Example : AtMostSeqCard$(2, 5, 8, [x_1, ..x_{22}])$

S	1	0	1	0	0	.	.	0	0	0	1	1	0	0	0	0	1	0	0	0	0	1
$\texttt{leftmost}(S(x_i))$	1	0	1	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	1
$Left[i]$	1	1	2	2	2	3	3	3	3	3	4	5	5	5	5	5	6	6	6	6	6	7

$Left[22] = 7 < 8$: FAILURE

$max(i)$	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1
$S^{\cdot *}$	1	.	1	.	.	.	.	.	.	.	1	1	.	.	.	0	.	0	0	0	0	.

The final explanation size $|S^{\cdot *}|$ is 9 while the naive one ($|S|$) is 20.

Explaining pruning

explanation for $x \leftarrow k$?

1. Add $x \not\leftarrow k$ to the instantiation where the pruning was performed.
2. Use the previous procedure to explain the failure on the new instantiation.

PB & SAT Modelling

Variables:

- c_i^j : c_i^j is *true* iff the class of the ith slot is j.
- y_i^j : y_i^j is *true* iff the ith vehicle requires option j.

Constraints:

- Demand constraints : $\forall j \in [1..k], \sum_i c_i^j = D_j$
- Capacity constraints : $\sum_{l=i}^{i+q_j-1} y_l^j \leq u_j$
- Channelling:
  - $\forall i \in [1..n], \forall l \in [1..k]$, we have :
    - $\forall j \in \mathcal{O}_l, \overline{c_i^l} \vee y_i^j$
    - $\forall j \notin \mathcal{O}_l, \overline{c_i^l} \vee \overline{y_i^j}$
  - a redundant clause : $\forall i \in [1..n], j \in [1..m], \overline{y_i^j} \vee \bigvee_{l \in \mathcal{C}_j} c_i^l$
- $\forall i \in [1..n], \sum_j c_i^j = 1$

PB & SAT Modelling

Variables:

- c_i^j : c_i^j is *true* iff the class of the ith slot is j.
- y_i^j : y_i^j is *true* iff the ith vehicle requires option j.

Constraints:

- Demand constraints : $\forall j \in [1..k], \sum_i c_i^j = D_j$
- Capacity constraints : $\sum_{l=i}^{i+q_j-1} y_l^j \leq u_j$
- Channelling:
  - $\forall i \in [1..n], \forall l \in [1..k]$, we have :
    - $\forall j \in \mathcal{O}_l, \overline{c_i^l} \vee y_i^j$
    - $\forall j \notin \mathcal{O}_l, \overline{c_i^l} \vee \overline{y_i^j}$
  - a redundant clause : $\forall i \in [1..n], j \in [1..m], \overline{y_i^j} \vee \bigvee_{l \in \mathcal{C}_j} c_i^l$
- $\forall i \in [1..n], \sum_j c_i^j = 1$

SAT model? encode Cardinality constraints: Sequential counter, Cardinality Networks, Sorting network, etc.

PB & SAT Modelling

Variables:

- c_i^j : c_i^j is *true* iff the class of the ith slot is j.
- y_i^j : y_i^j is *true* iff the ith vehicle requires option j.

Constraints:

- Demand constraints : $\forall j \in [1..k], \sum_i c_i^j = D_j$
- Capacity constraints : $\sum_{l=i}^{i+q_j-1} y_l^j \leq u_j$
- Channelling:
  - $\forall i \in [1..n], \forall l \in [1..k]$, we have :
    - $\forall j \in \mathcal{O}_l, \overline{c_i^l} \vee y_i^j$
    - $\forall j \notin \mathcal{O}_l, \overline{c_i^l} \vee \overline{y_i^j}$
  - a redundant clause : $\forall i \in [1..n], j \in [1..m], \overline{y_i^j} \vee \bigvee_{l \in \mathcal{C}_j} c_i^l$
- $\forall i \in [1..n], \sum_j c_i^j = 1$

SAT model? encode Cardinality constraints: Sequential counter, Cardinality Networks, Sorting network, etc.

Sequential Counter (SC) [Sin05]

Encoding $\sum_{i \in [1..n]} x_i = d$ to a CNF ?

- Variables:
  - $s_{i,j}$: $\forall i \in [0..n], \forall j \in [0..d+1]$, $s_{i,j}$ is *true* iff $\sum_{k \in [1..i]} x_k \geq j$
- Encoding: $\forall i \in [1..n]$
  - Clauses for restrictions on the same level: $\forall j \in [0..d+1]$
    1. $\neg s_{i-1,j} \vee s_{i,j}$
    2. $x_i \vee \neg s_{i,j} \vee s_{i-1,j}$
  - Clauses for increasing the counter, $\forall j \in [1..d+1]$
    3. $\neg s_{i,j} \vee s_{i-1,j-1}$
    4. $\neg x_i \vee \neg s_{i-1,j-1} \vee s_{i,j}$
  - Initial values for the bounds of the counter:
    5. $s_{0,0} \wedge \neg s_{0,1} \wedge s_{n,d} \wedge \neg s_{n,d+1}$

Example $\sum_{i\in[1..8]} x_i = 2$

3	0	0	0	0	0	0	0	0	0
2	0	0	.	.	.	.	.	.	1
1	0	.	.	.	.	.	.	1	1
0	1	1	1	1	1	1	1	1	1
$s_{i,j}$	0	1	2	3	4	5	6	7	8
x_i		.	.	.	.	.	.	.	.

Example $\sum_{i \in [1..8]} x_i = 2$

3	0	0	0	0	0	0	0	0	0
2	0	0	.	.	.	.	.	.	1
1	0	.	.	.	.	.	.	1	1
0	1	1	1	1	1	1	1	1	1
$s_{i,j}$	0	1	2	3	4	5	6	7	8
x_i		.	.	.	.	.	.	.	.

3	0	0	0	0	0	0	0	0	0
2	0	0	.	.	.	.	.	.	1
1	0	.	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1	1	1
$s_{i,j}$	0	1	2	3	4	5	6	7	8
x_i		.	1	.	.	.	.	.	.

Example $\sum_{i \in [1..8]} x_i = 2$

3	0	0	0	0	0	0	0	0	0
2	0	0	.	.	.	.	.	.	1
1	0	.	.	.	.	.	.	1	1
0	1	1	1	1	1	1	1	1	1
$s_{i,j}$	0	1	2	3	4	5	6	7	8
x_i		.	.	.	.	.	.	.	.

3	0	0	0	0	0	0	0	0	0
2	0	0	.	.	.	.	.	.	1
1	0	.	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1	1	1
$s_{i,j}$	0	1	2	3	4	5	6	7	8
x_i		.	1	.	.	.	.	.	.

3	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	1	1
1	0	0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1	1	1
$s_{i,j}$	0	1	2	3	4	5	6	7	8
x_i		0	1	0	0	0	0	1	0

Example $\sum_{i\in[1..8]} x_i = 2$

3	0	0	0	0	0	0	0	0	0
2	0	0	.	.	.	.	.	.	1
1	0	.	.	.	.	.	.	1	1
0	1	1	1	1	1	1	1	1	1
$s_{i,j}$	0	1	2	3	4	5	6	7	8
x_i		.	.	.	.	.	.	.	.

3	0	0	0	0	0	0	0	0	0
2	0	0	.	.	.	.	.	.	1
1	0	.	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1	1	1
$s_{i,j}$	0	1	2	3	4	5	6	7	8
x_i		.	1	.	.	.	.	.	.

3	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	1	1
1	0	0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1	1	1
$s_{i,j}$	0	1	2	3	4	5	6	7	8
x_i		0	1	0	0	0	0	1	0

$\rightarrow$Unit Propagation on the SC encoding enforces AC on the cardinality constraint $\sum_{i\in[1...n]} x_i = d$.

Extension to AtMostSeqCard

AtMostSeqCard: AtMost⊕ Cardinality→[SC] !

Extension to AtMostSeqCard

AtMostSeqCard: AtMost⊕ Cardinality→[SC] !

Using a similar encoding of the Gen-Sequence constraint [Bac07, BNQ+07] [SCS].

7. $\neg s_{i,j} \vee s_{i-q,j-u}$

Extension to AtMostSeqCard

AtMostSeqCard: AtMost⊕ Cardinality→[SC] !

Using a similar encoding of the Gen-Sequence constraint [Bac07, BNQ+07] [SCS].

8 $\neg s_{i,j} \vee s_{i-q,j-u}$

Proposition

The level of pruning using (SCS) is incomparable with SC on each AtMost(SCA).

Configuration

- SAT :
  1. *SAT (1)* SC
  2. *SAT (2)* SCS
  3. *SAT (3)* SC $\oplus$ SCS.
- Mistral as a hybrid CP/SAT solver
  1. *Hybrid (VSIDS)*
  2. *Hybrid (Slot)*
  3. *Hybrid (Slot $\rightarrow$ VSIDS)*
  4. *Hybrid (VSIDS $\rightarrow$ Slot)*
- *pseudo Boolean*: MiniSat+
- *CP*: [Mistral]

Table: Evaluation of the models

Method	sat[easy] (74 × 5)			sat[hard] (7 × 5)			unsat/unknown (28 × 5)		
	#suc	avg fails	time	#suc	avg fails	time	#suc	avg fails	time
SAT (1)	370	2073	1.71	28	337194	282.35	**85**	**249301**	**105.07**
SAT (2)	370	1077	1.18	30	42790	33.02	67	217103	182.23
SAT (3)	370	667	1.30	35	50233	66.23	74	137639	70.47
Hybrid (VSIDS)	370	903	0.23	16	207211	286.32	35	177806	224.78
Hybrid (VSIDS → Slot)	370	739	0.23	35	76256	64.52	37	204858	248.24
Hybrid (Slot → VSIDS)	370	132	0.04	34	4568	2.50	37	234800	287.61
Hybrid (Slot)	370	132	0.04	**35**	**6304**	**3.75**	23	174097	299.24
CP	**370**	**43.06**	**0.03**	35	57966	16.25	0	-	-
pseudo Boolean	277	538743	236.94	0	-	-	43	175990	106.92

Table: Evaluation of the models

Method	sat[easy] (74 × 5)			sat[hard] (7 × 5)			unsat/unknown (28 × 5)		
	#suc	avg fails	time	#suc	avg fails	time	#suc	avg fails	time
SAT (1)	370	2073	1.71	28	337194	282.35	**85**	**249301**	**105.07**
SAT (2)	370	1077	1.18	30	42790	33.02	67	217103	182.23
SAT (3)	370	667	1.30	35	50233	66.23	74	137639	70.47
Hybrid (VSIDS)	370	903	0.23	16	207211	286.32	35	177806	224.78
Hybrid (VSIDS → Slot)	370	739	0.23	35	76256	64.52	37	204858	248.24
Hybrid (Slot → VSIDS)	370	132	0.04	34	4568	2.50	37	234800	287.61
Hybrid (Slot)	370	132	0.04	**35**	**6304**	**3.75**	23	174097	299.24
CP	**370**	**43.06**	**0.03**	35	57966	16.25	0	-	-
pseudo Boolean	277	538743	236.94	0	-	-	43	175990	106.92

- Finding solutions quickly :
  - CP-based models are difficult to outperform!
  - Overall, the best method on satisfiable instances is the hybrid solver using a pure CP heuristic.
  - with *VSIDS*, MiniSat on the strongest encodings has good results!
  - Propagation is very important to find solutions quickly when they exist, by keeping the search “on track” and avoiding exploring large unsatisfiable subtrees.

Table: Evaluation of the models

Method	sat[easy] (74 × 5) #suc	avg fails	time	sat[hard] (7 × 5) #suc	avg fails	time	unsat/unknown (28 × 5) #suc	avg fails	time
SAT (1)	370	2073	1.71	28	337194	282.35	**85**	**249301**	**105.07**
SAT (2)	370	1077	1.18	30	42790	33.02	67	217103	182.23
SAT (3)	370	667	1.30	35	50233	66.23	74	137639	70.47
Hybrid (VSIDS)	370	903	0.23	16	207211	286.32	35	177806	224.78
Hybrid (VSIDS → Slot)	370	739	0.23	35	76256	64.52	37	204858	248.24
Hybrid (Slot → VSIDS)	370	132	0.04	34	4568	2.50	37	234800	287.61
Hybrid (Slot)	370	132	0.04	**35**	**6304**	**3.75**	23	174097	299.24
CP	**370**	**43.06**	**0.03**	35	57966	16.25	0	-	-
pseudo Boolean	277	538743	236.94	0	-	-	43	175990	106.92

- For proving unsatisfiability
  - Clause learning is by far the most critical factor.
  - Surprisingly, the “lightest” encoding gave best results!

Conclusion & Future research

Contributions

- First non-trivial SAT encoding for the car-sequencing problem.
- A linear time explanation for the AtMostSeqCard constraint
- Closing 13 out of the 23 large open instances.

Future research

- Can we generate optimal explanations for AtMostSeqCard?
- Other SAT-encoding for the Cardinality constraint?
- Optimisation problems?

Thank you!

Fahiem Bacchus.
GAC Via Unit Propagation.
In *Proceedings of CP*, pages 133–147, 2007.

Sebastian Brand, Nina Narodytska, Claude-Guy Quimper, Peter J. Stuckey, and Toby Walsh.
Encodings of the Sequence Constraint.
In *Proceedings of CP*, pages 210–224, 2007.

Carsten Sinz.
Towards an Optimal CNF Encoding of Boolean Cardinality Constraints.
In *Proceedings of CP*, pages 827–831, 2005.